[DETAILED DESCRIPTION OF THE INVENTION]

[Industrial Applicability]

The present invention relates to a method of object image recognition processing in which a plurality of identification target categories are defined on the basis of a difference in the shape and the surface texture, and in which positioning of the comparison pattern of a target object is carried out in order to identify the class to which the object belongs on the basis of the image information of the object which has certain common characteristics to other objects in the shape and the surface texture.

In the following description, the face of a person is used as an example of the identification target object having the above-mentioned characteristics, whereby an exemplary application of the invention to person identification by face image is described below. However, it should be noted that the invention is not restricted to this, but is applicable widely to image recognition of various objects having the above-mentioned characteristics.

[Prior Art]

In order to identify the class to which an object belongs using image information, it is necessary to position accurately a region (referred to as a comparison pattern) in which comparison with standard patterns prepared for respective categories in advance is carried out within an input image.

In the image recognition of an object, such as the face of a person, in which the difference of the identification target class can be defined by a difference in the shape and the surface texture, whereas there are still class-independent common characteristics in the shape and the surface texture, a method has been proposed in which the positioning of the comparison pattern is carried out using, as the reference, several characteristic points expected to be extracted from the input image regardless of the difference in the target object class.

For example, in the recognition of the face image of a person, a method in which the right and left eyes and the lips which are the representative points in the region of the features of the face are used as the reference points in the positioning of the comparison pattern is reported in Reference [1] (Akamatsu, Sasaki, and Suenaga, "Basic discussion of person recognition by front face image," Television Society Technical Report VAI90-36, October 1990).

In general, in image processing, it is very difficult to extract automatically and uniquely such characteristic points corresponding correctly to the features of the face from the face image. In order to treat this difficulty, a large number of possible characteristic points are first extracted so as to avoid the loss of the characteristic points corresponding correctly to the features. Then, a plurality of candidates of comparison patterns are obtained by cutting out each combination

of characteristic points which could be used as the reference points in cutting out among the possible characteristic points. After that, a testing process is carried out using a standard template prepared for the comparison pattern to be extracted from the correct characteristic points. Finally, a correct comparison pattern is selected among the candidates. This method is described in Reference [2] ("Method of face reference point extraction," Japanese Patent Unexamined Application No. Hei 3-144540).

[Problems to be Solved by the Invention]

In the method described in Reference [2], in the testing process for selecting a correct normalized pattern among the comparison pattern candidates, solely used was a standard template of a kind which represents a sample set of the comparison pattern obtained when the reference points are extracted correctly. Nevertheless, in this method, in case of an input face image of an unknown person which has not been used in the generation of the template, or alternatively in case that the illumination condition has been changed, the distance increases between the comparison pattern cut out in the combination of the characteristic points corresponding correctly to the reference points and the standard template. This reduces the difference in the distance in comparison with the case of the distance between the comparison pattern cut out in a combination of characteristic points corresponding incorrectly to the

reference points and the standard template. This increases the possibility that a comparison pattern corresponding to an incorrect normalized condition is selected. Accordingly, there has been the problem that the reliability in the positioning of the comparison pattern is reduced.

The invention has been devised in order to resolve such a problem. An object of the invention is to provide a technique capable of carrying out more stably the positioning of an identification target object by using, as the reference, the characteristics extracted from an image.

[Means for Solving the Problem]

In the positioning process in the invention, at first, characteristic points serving as the candidates of reference points for positioning are extracted by characteristic point extraction from an input pattern. Then, pattern positioning is carried out for each combination of reference point candidates, whereby a plurality of comparison pattern candidates are obtained each by cutting out a part of the image. Each comparison pattern candidate is compared with a plurality of templates each of which corresponds to each positioning condition and is stored in a positioning template dictionary in advance. On the basis of the result of the comparison process, a comparison pattern which is determined to be normalized correctly in the position, the size, and the like is selected and output among the comparison pattern candidates. This approach characterizes the invention.

In the above-mentioned positioning template dictionary, regarding a large number of images prepared in advance as learning samples of identification target objects, the set of cut-out images obtained by normalization process on a large number of combinations of points capable of serving as cut-out reference points among the extracted characteristic points is divided into pattern classes each having a similar structure. The Karhunen-Loeve expansion method is applied onto the set of sample images in each class, whereby the average image and the orthonormal base image group which describe the subspace satisfied by the sample image set are obtained, and then registered as the template for the class into the dictionary. On the other hand, in the process of comparison between an arbitrary comparison pattern candidate obtained from the input pattern and the template, the distance value between the comparison pattern candidate and the subspace of the class is calculated using the average image and the orthonormal base images. Further, the process of selecting a comparison pattern to be used for identification process among the comparison pattern candidates for the input pattern is carried out on the basis of the distance value between each comparison pattern candidate and the subspace of each class. This approach also characterizes the invention.

[Operation of the Invention]

According to the above-mentioned means, in the image

recognition of an object having common characteristics in the shape or the surface texture regardless of class difference, when the positioning of the object is carried out using, as the reference points, the characteristic points extracted from an input image, parameters describing the subspace formed by each pattern set are separately registered as the standard template used for selecting one satisfying the correct normalization condition among the comparison pattern candidates generated in a plural number, not only for sample sets of cut-out patterns satisfying the correct normalization condition, but also sample sets having higher occurrence frequency among the cut-out results by combinations of characteristic points causing an incorrect normalization condition. Then, the distance is compared with each other between the respective comparison pattern candidates obtained from the input image and the respective standard templates. By virtue of this, even in case of the input of an unknown object which has not been used in the generation of the standard templates, or alternatively in case that the illumination condition has been changed in the input image, the positioning of the identification target object can be carried out more stably.

[Embodiments of the Invention]

An embodiment of the invention is described below with reference to the drawings. FIGURE 1 is a block diagram showing the functional configuration of an image recognition system used

for the description of an embodiment of a method of image recognition of an object according to the invention.

In FIGURE 1, numeral 1 indicates an arbitrary identification target object. Numeral 2 indicates an image input process section. Numeral 3 indicates a comparison pattern positioning process section. Numeral 4 indicates a characteristics extraction process section. Numeral 5 indicates a category identification dictionary. Numeral 6 indicates a standard characteristic pattern comparison process section. Numeral 7 indicates an identification determination process section. Numeral 8 indicates an identification result. The comparison pattern positioning process section 3 comprises a positioning reference point candidate extraction section 31, a comparison pattern candidate cut-out process section 32, a comparison pattern candidate storage section 33, a positioning template dictionary 34, a positioning template comparison process section 35, and a comparison pattern selection process section 36.

The operation of image recognition process of an object is described below. Using the image input process section 2 composed of a television camera or the like, two-dimensional image of an identification target object 1 is input. The comparison pattern positioning process section 3 carries out the detection of a target object in the input image and the normalization process with respect to the position, the size,

and the like, and thereby obtaining a comparison pattern used as the object of the subsequent identification process. The characteristics extraction process section 4 extracts a characteristic pattern represented mathematically by an multidimensional vector, from the comparison pattern. The standard characteristic pattern comparison process section 6 compares the characteristic pattern of the input image with the standard characteristic pattern of each category stored in the category identification dictionary 5 in advance. The identification determination process section 7 determines the category to which the identification target object belongs, on the basis of the comparison result.

As the detailed processes carried out in the characteristics extraction process section 4, the standard characteristic pattern comparison process section 6, and the identification determination process section 7, various methods have been proposed depending on the type of identification target object and the type of obtained input image. For example, an embodiment in which the face of a person is used as the identification object and in which two-dimensional density image thereof is input is described in Reference [3] (Akamatsu, Sasaki, Fukamachi, and Suenaga, "Evaluation of the application of pattern description method by KL expansion to face image identification," Singaku-Gihou PRU90-52, March 1991).

An embodiment in which the face of a person is used as

the identification object in the process of the comparison pattern positioning process section 3 is described below. In the selection of the comparison pattern for the input face image, similarly to Reference [1], affine transformation is applied to the original image such that the characteristic points (Er, El, and M) corresponding to the region centers of the right and left eyes and the lips, respectively, satisfy a certain spatial positional relation (defined by a parameter D) as shown in FIGURE 2. After that, a 128×128 density image is obtained by sampling a square region defined by certain parameters C1-C4 with respect to those reference points. The following description is made for this case. FIGURE 2 shows the situation in which a comparison region is extracted from a face image shown in the left portion of the figure.

The positioning reference point candidate extraction section 31 first extracts the candidates of characteristic points corresponding to the face internal features of the right and left eyes and the lips to be used as the reference points in the cutting-out of the comparison pattern from the input face image. In an embodiment of this process, with focusing attention on the color characteristics of such face internal features, a method of region division process of a color face image described in Reference [2] is applicable. Nevertheless, in the present step, it is very difficult to obtain stably and uniquely the reference points corresponding correctly to the features. Thus,

in general, when the extraction of the correct region is to be secured, a plurality of incorrect reference point candidates are unavoidably obtained at the same time.

Accordingly, in the comparison pattern candidate cut-out process section 32, the number of candidates of possible combination of the correct reference points among a plurality of combinations of characteristic points obtained by the positioning reference point candidate extraction section 31 is reduced on the basis of an ad hoc determination process on the color and the positional relation of the regions as described in Reference [2]. After that, with the assumption that each candidate is the combination of the correct reference points, each comparison pattern candidate is cut out in the method shown in FIGURE 2, and then stored in the comparison pattern candidate storage section 33.

Here, it is assumed that K comparison pattern candidates are obtained for an input face image. The density value of each density image is expressed by each one-dimensional vector X1, X2, ..., XK. The positioning template comparison process section 35 and the comparison pattern selection process section 36 compare each of the K comparison pattern candidates X1, X2, ..., XK with each template corresponding to each class stored in the positioning template dictionary 34, and thereby select a comparison pattern determined as the nearest to the correct cut-out result in the sense of FIGURE 2.

In the invention, it is expected that the common structural characteristics of the face objects would result in a certain trend in the distribution of the occurrence positions of the characteristic points extracted in the input image by the positioning reference point candidate extraction section 31, in other words, in the structure of the comparison pattern cut out from the input image by the comparison pattern candidate cut-out process section 32 using the characteristic points as the reference. On the basis of this expectation, a large number of comparison pattern candidates obtained from the learning samples are classified into several classes in advance (also included is one corresponding to the normalized pattern obtained from the correct reference points as shown in FIGURE 2). This approach characterizes the invention.

FIGURE 3 is a conceptual diagram showing an example of classification of the comparison pattern candidates. In the figure, Class 1 indicates the correct normalization as shown in FIGURE 2. Class 2 indicates that the size is correct, but that the image is rotated clockwise. Class 3 indicates that the size is correct, but that the image is rotated counterclockwise. Class 4 indicates that the size is reduced, but that the position is approximately correct. Class 5 indicates that the size is reduced, but that the image is rotated clockwise. Class 6 indicates that the size is reduced, but that the image is rotated counterclockwise. Class 7 indicates that the size is enlarged,

but that the position is approximately correct. Class 8 indicates that the size is enlarged, but that the image is rotated clockwise. Class 9 indicates that the size is enlarged, but that the image is rotated counterclockwise.

Such classification of the cut-out results is thought to be based on the face-image specific characteristics which are easily detected incorrectly as the characteristic points representing the right and left eyes and the lips used as the reference points, that is, on the common characteristics of faces such as the presence of the eyebrows and the nasal cavity. Obviously, there are a large number of other possible classifications actually. A template is prepared for each class using the learning samples classified into the class.

An example of the template for each class is the average pattern of the learning samples belonging to the class as described in Reference [2]. However, in the present embodiment, the following method based on the framework of subspace method is introduced in order to permit class determination with a more precise cut-out condition by comparison with the template.

It is assumed that there are M comparison patterns of learning samples classified into the same cut-out condition. The N-dimensional characteristic vectors each having the density value thereof are denoted by X1, X2, ..., Xm, ..., XM. In this case, the template of this cut-out class is obtained as follows.

(1) Average vector μ over the M characteristic vectors is

obtained.

(2) Orthonormal base image sequence U1, U2, ..., UL of the subspace of the M sample sets is obtained by the Karhunen-Loeve expansion of the differential image set {Xm−μ} of the sample image.

Here, the orthonormal base image sequence U1, U2, ..., UL is obtained as the eigenvecotrs corresponding to the upper L engenvalues of the sample covariance matrix R of the sample image sets. The detailed method of this calculation is described in Reference [4] (Erlicki Yao, "Pattern recognition and subspace method," Sangyo Tosho). Hence, the description is omitted herein.

The average vector μ calculated from the comparison pattern sets as described above and the L orthonormal base image sequence U1, U2, ..., UL to be used as templates are stored in the positioning template dictionary 34 for each class classified depending on the exemplary cut-out condition shown in FIGURE 3.

The positioning template comparison process section 35 compares the K comparison pattern candidates X1, X2, ..., XK stored in the comparison pattern candidate storage section 33 with the template of each class stored in the positioning template dictionary 34. Here, the density vector representation of an arbitrary comparison pattern candidate is denoted by X. The template of the compared class is composed of the average vector μ and the L orthonormal base image sequence U1, U2, ..., UL. Then, in an embodiment of the comparison process, the value d

obtained by the following formula is used as the comparison result between the comparison pattern candidate X and the class.

$$d^2 = (X - \mu)^t(X - \mu) - \sum_{i=1}^{L} \{(X - \mu)^t U i\}^2 \qquad \cdots (1)$$

The value d corresponds to the Euclid distance from a point X in the N-dimensional characteristic space to the subspace of the class defined by the average vector μ and the L orthonormal base image sequence U1, U2, ..., UL.

Let K be the number of comparison pattern candidates, while let C be the number of cut-out condition classes each provided with a template. Then, the result obtained for an arbitrary input pattern by the positioning template comparison process section 35 is a K×C array, that is, a comparison pattern-template comparison table, as shown in FIGURE 4. Here, the comparison result (the value d obtained according to the Formula (1)) of the k-th comparison pattern candidate with the subspace of the c-th cut-out condition category is denoted by d_{kc}.

Using the comparison pattern-template comparison table, the comparison pattern selection process section 36 selects a comparison pattern determined as the nearest to the correct cut-out result in the sense of FIGURE 2 among the comparison pattern candidates. FIGURE 5 shows an embodiment of this determination process. Here, in the example of the comparison pattern-template comparison table shown in FIGURE 4, it is assumed that Class 1 corresponding to the first row represents

the comparison pattern corresponding to the correct cut-out condition shown in FIGURE 2.

(ST1) The minimum value of d_{kc} for an arbitrary comparison pattern candidate k is denoted by min1(k). The cut-out condition category c giving this is denoted by Cmin1(k). The second minimum value is denoted by min2(k).

(ST2) It is checked whether there is a comparison pattern candidate k in which Cmin1(k) is 1. In case of yes, the procedure goes to ST3. In case of no, the procedure goes to ST8.

(ST3) It is checked whether the candidate satisfying the above-mentioned relation is the only one or not. In case of yes, the procedure goes to ST4. In case of no, the procedure goes to ST5.

(ST4) The comparison pattern candidate is selected.

(ST5) The plural comparison pattern candidates are denoted by K1, K2,...

(ST6) It is checked whether there is a sufficiently reliable comparison pattern candidate Kn which satisfies that min2(Kn) —min1(Kn) > 0. In case of yes, the procedure goes to ST7. In case of no, the procedure goes to ST8.

(ST7) Among the Kn's which satisfy the above-mentioned relation, a comparison pattern candidate corresponding to the minimum Kn having the minimum min1(k) is selected.

(ST8) The result is rejected.

(ST9) Retrial.

In the prior art, in case of an input face image of an unknown person which has not been used in the generation of the standard template, or alternatively in case that the illumination condition has been changed, the comparison value with the standard template of Class 1 has been increased even for the comparison pattern to be determined as correct normalization condition in the sense of FIGURE 2. This has caused the cases in which the correct comparison pattern can not be selected only on the basis of the comparison of the values in the first row (Class 1) in the comparison pattern-template comparison table. However, in the invention, by virtue of the process flow shown in FIGURE 5 in which the determination process is carried out also for the values of comparison with other classes corresponding to typical incorrect cut-out conditions, the reliability is improved in the positioning of the comparison pattern.

FIGURE 6 schematically shows the effect of the present process. In the figure, the left cluster indicates the subspace of Class 1, that is, of the comparison pattern set corresponding to the correct normalization condition in the sense of FIGURE 2, within the N-dimensional characteristic space. The right cluster indicates the subspace of an arbitrary Class c, that is, of the comparison pattern set corresponding to an incorrect normalization condition. Points a and b indicate two comparison pattern candidates for an input pattern, within the N-dimensional

characteristic space. It is assumed that the point a satisfies the correct normalization condition, but that the point b does not satisfy the correct normalization condition. The point a lies at a position away from the subspace of Class 1 because of the above-mentioned reason. And hence, it is assumed that

$$d_{a1} > d_{b1}$$

In this case, when the nearness to the subspace of Class 1 is solely considered, the comparison pattern candidate b may be selected instead of the comparison pattern candidate a. However, the subspace nearest to the point a is the Class 1 subspace. In contrast, the point c is nearer to the Class c subspace than to the Class 1 subspace. That is,

$$d_{a1} < d_{ac} \qquad \text{for all c}$$

$$d_{b1} > d_{bc}$$

Accordingly, the point a is selected instead of the point b, according to the process flow shown in FIGURE 5.

The invention has been described above in detail for an embodiment of an image recognition system in which the face of a person is used as the identification target. However, the invention is not restricted to this. Various modifications are obviously possible without departing from the scope of the invention.

[Effects of the Invention]

As described above, according to the invention, among the characteristic points extracted from the input image as the

reference point candidates used for the positioning of the identification target, parameters of subspaces of the sets of comparison patterns assigned to respective classes are prepared as the plural standard templates in the positioning template dictionary, not only for sample sets of cut-out patterns satisfying the correct normalization condition, but also sample sets having higher occurrence frequency among the cut-out results by combinations of characteristic points causing an incorrect normalization condition.

By virtue of this, by calculating the distance from the subspace corresponding to each cut-out condition class, each comparison pattern candidate obtained from the input image can be evaluated in the likelihood of the cut-out condition class to which the comparison pattern candidate belongs. By virtue of this, even in case of the input of an unknown object not included in the learning samples having been used in the generation of the standard templates, or alternatively in case that the illumination condition has been substantially changed in the input image, the positioning of the comparison pattern to the identification target object can be carried out more stably and more accurately. This improves the identification precision in the image recognition system.